The first step in sugar transport: crystal structure of the amino terminal domain of enzyme I of the *E. coli* PEP: sugar phosphotransferase system and a model of the phosphotransfer complex with HPr

D-I Liao[1], E Silverton[1], Y-J Seok[2], BR Lee[2†], A Peterkofsky[2] and DR Davies[1*]

Background: The bacterial phosphoenolpyruvate (PEP): sugar phosphotransferase system (PTS) transports exogenous hexose sugars through the membrane and tightly couples transport with phosphoryl transfer from PEP to the sugar via several phosphoprotein intermediates. The phosphate group is first transferred to enzyme I, second to the histidine-containing phosphocarrier protein HPr, and then to one of a number of sugar-specific enzymes II. The structures of several HPrs and enzymes IIA are known. Here we report the structure of the N-terminal half of enzyme I from *Escherichia coli* (EIN).

Results: The crystal structure of EIN (MW ~30 kDa) has been determined and refined at 2.5 Å resolution. It has two distinct structural subdomains; one contains four α helices arranged as two hairpins in a claw-like conformation. The other consists of a β sandwich containing a three-stranded antiparallel β sheet and a four-stranded parallel β sheet, together with three short α helices. Plausible models of complexes between EIN and HPr can be made without assuming major structural changes in either protein.

Conclusions: The α/β subdomain of EIN is topologically similar to the phosphohistidine domain of the enzyme pyruvate phosphate dikinase, which is phosphorylated by PEP on a histidyl residue but does not interact with HPr. It is therefore likely that features of this subdomain are important in the autophosphorylation of enzyme I. The helical subdomain of EIN is not found in pyruvate phosphate dikinase; this subdomain is therefore more likely to be involved in phosphoryl transfer to HPr.

Addresses: [1]Laboratory of Molecular Biology, NIDDK, National Institutes of Health, Bethesda, MD 20892 and [2]Laboratory of Biochemical Genetics, NHLBI, National Institutes of Health, Bethesda, MD 20892, USA.

[†]Present address: Department of Biological Education, Seo-Won University, Mochung-Dong, Chong-Ju City, Chung-Buk, South Korea.

*Corresponding author.
E-mail: drd@vger.niddk.nih.gov

Key words: enzyme I, HPr, phosphotransferase, sugar transport, X-ray crystallography

Received: 1 April 1996
Revisions requested: 26 April 1996
Revisions received: 28 May 1996
Accepted: 4 June 1996

Structure 15 July 1996, 4:861–872

Introduction

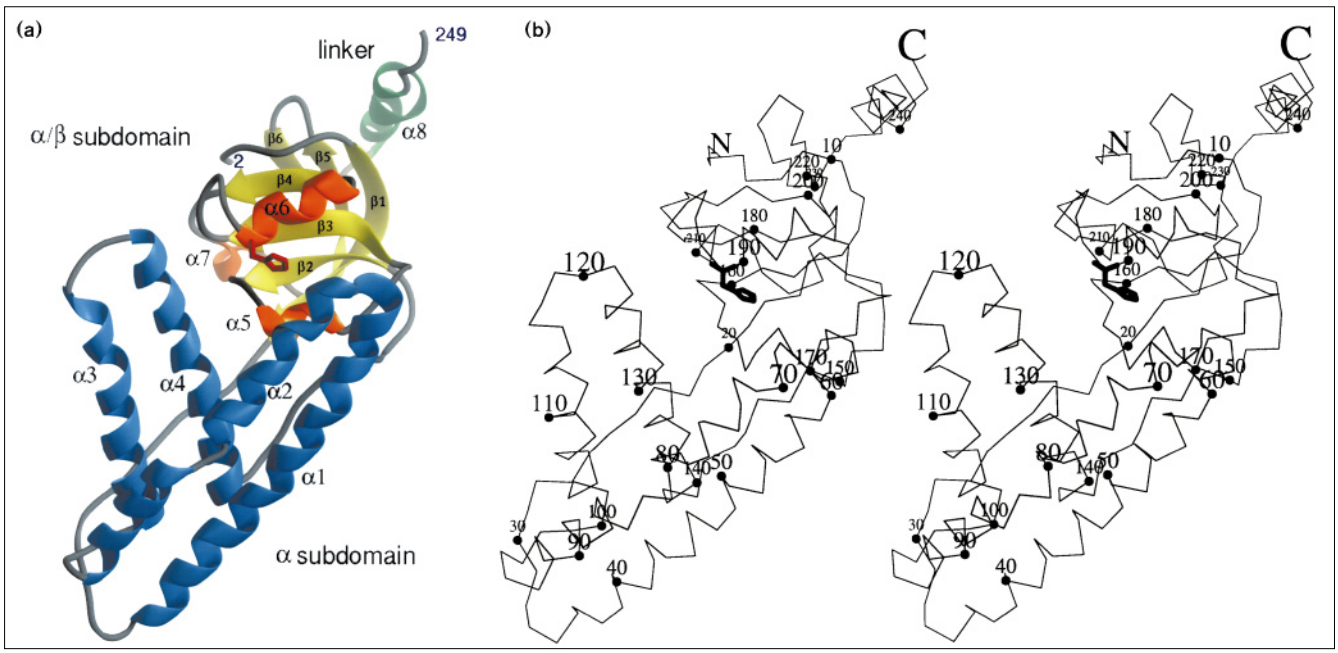
The phosphoenolpyruvate:sugar phosphotransferase system (PTS) [1], is capable of effecting the coupled phosphorylation and translocation of numerous sugars across the cytoplasmic membrane; the system is found throughout the bacterial kingdom. The multiprotein PTS includes a general protein, enzyme I (EI), which is autophosphorylated by phosphoenolpyruvate (PEP) on the Nε atom of a conserved histidine residue (His189 in *Escherichia coli*). Phosphorylated EI acts as a phosphoryl donor to another general protein, the low molecular weight HPr, which is phosphorylated on the Nδ atom of a histidine residue (His15 in *E. coli*). Phosphorylated HPr interacts with numerous membrane associated proteins, termed enzymes II, which are responsible for the sugar-specific translocation/phosphorylation reactions.

HPr (~9 kDa) and the ~18 kDa enzyme IIA^{glc} ($EIIA^{glc}$), a membrane associated protein specific for glucose transport, have been extensively characterized at the structural level. Both NMR and crystallographic results have been used to describe precise structures for these proteins. For a complete review, see [2].

In contrast to the wealth of structural understanding for HPr and $EIIA^{glc}$, there is a paucity of information regarding the structure of EI. A combination of proteolysis [3] and thermodynamic [4] studies suggested that EI is composed of a proteolytically stable N-terminal domain (EIN) and a proteolytically unstable C-terminal domain (EIC), connected to EIN by a linker. Each domain comprises about 50% of the molecule. EIN cannot be autophosphorylated by PEP, but can participate in reversible phosphoryl transfer with phospho–HPr (P–HPr) [4,5]. It has therefore been concluded, that the C-terminal domain of EI is necessary for the interaction with PEP or pyruvate. Studies of the C-terminal domain of EI using site-directed mutagenesis support the notion that it is important for interaction with PEP [6]. It appears that the C-terminal domain also plays a role in interaction with HPr; P–EI can only transfer its phosphoryl group to *E. coli* HPr, whereas P–EIN can phosphorylate HPrs from a variety of species [5].

The large size of EI (approximately 64 kDa) prohibits attempts to deduce its structure by NMR. Efforts in this and other laboratories to produce crystals of EI suitable for high-resolution crystallographic analysis have been unsuccessful.

Figure 1



The structure of EIN. (a) Ribbon representation of the EIN structure. His189 (shown in red) is located at the interface between the two domains. Helices in the α domain are blue and those in the α/β domain are orange. The linker is shown in green and the β strands in yellow, α helices and β sheets are numbered accordingly. (b) Stereo drawing showing the Cα backbone of the EIN molecule. Every tenth Cα atom is highlighted and labeled. The N and C termini are marked.

Consequently, attention has been directed to the analysis of EIN at the structural level. Highly purified preparations of a recombinant form of *E. coli* EIN [5] were found to crystallize readily. This report presents a three-dimensional structure of EIN based on crystallographic analysis. The data represent the first insight into the fold of EIN as well as to the mode of interaction of the protein with HPr.

Results and discussion

Structure of the protein

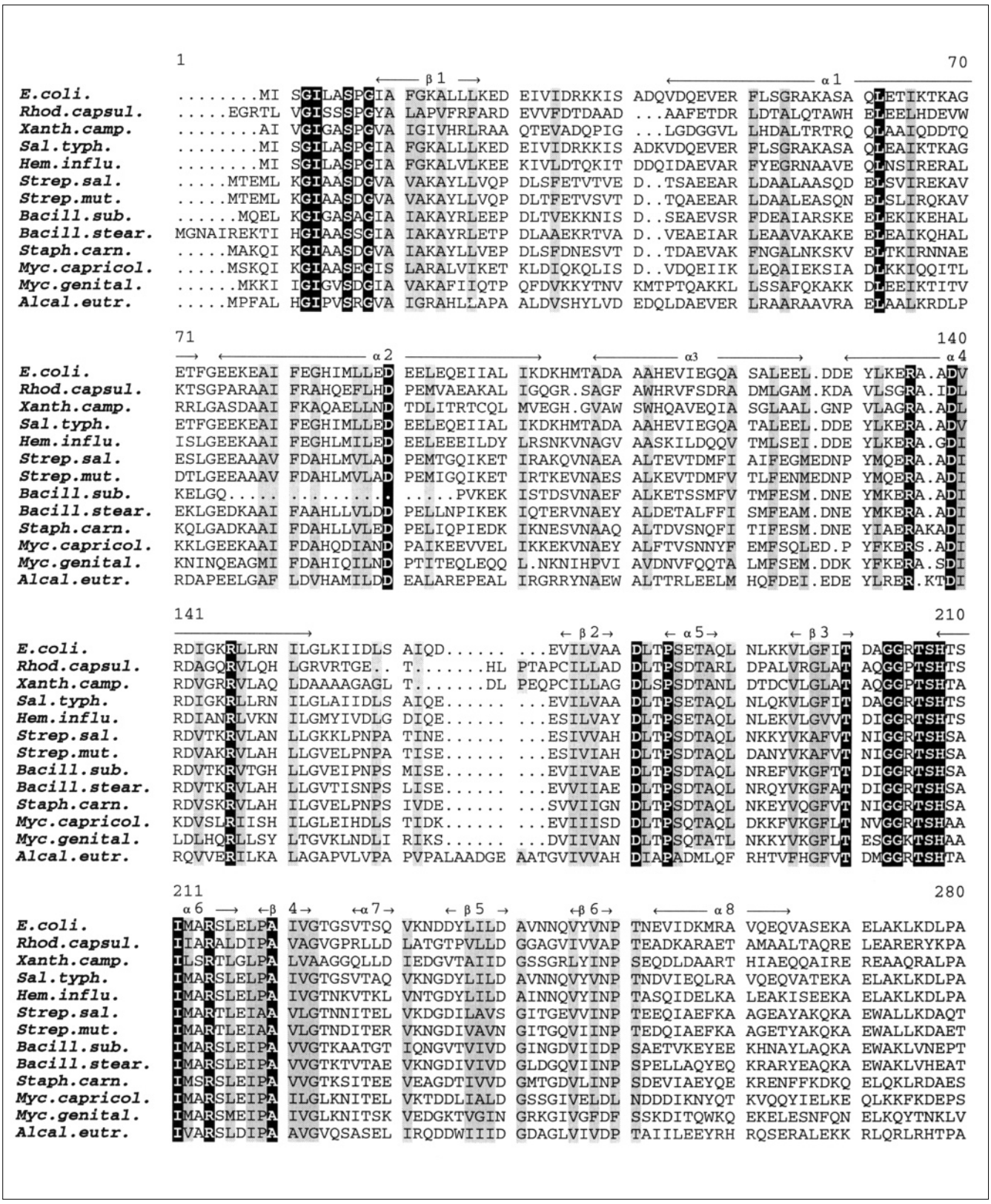
The model of EIN presented here contains residues 2–249 of the 259 residue polypeptide. The protein has an elongated two-subdomain structure (Fig. 1a,b), is about 66 Å long and its width varies from 16 Å to 26 Å, depending on the section of the molecule examined. The protein consists of an α subdomain, including α helices 1–4 (residues 30–142), and an α/β subdomain (residues 2–19 [β1], 156–229 [α5–7 and β2–6]). The α/β subdomain contains the phosphorylation site (His189). The two subdomains are connected by two extended loops (residues 20–29 and 143–155). In addition, there is a helical region (residues 230–248, α8) which serves as a linker to the C-terminal domain of E1 (not shown in the figure).

The α subdomain consists of two long helix hairpins packed in a claw-like conformation. The angle between the two hairpins is 60°. There is a large kink, corresponding to residues 81–83, between the fourth and fifth turns of the α2 helix. In the middle of the kink, the carboxyl group of the Asp82 side chain forms two hydrogen bonds to the main-chain nitrogen atoms of the next helical turn (residues 84 and 85). Asp82 is conserved in all EIs with known sequences except in *Bacillus subtilis* (see Fig. 2). Most of the sequence corresponding to the α2 helix is

Figure 2 (facing page)

Alignment of the N-terminal regions of EI proteins from 13 different organisms: *Escherichia coli* (*E. coli*, Genbank M10425); *Rhodobacter capsulatus* (*Rhod. capsul.*, SwissProt P23388); *Xanthomonas campestris* (*Xanth. camp.*, Genbank Z37113); *Salmonella typhimurium* (*Sal. typh.*, SwissProt P12654); *Hemophilus influenzae* (*Hem. influ.*, Genbank L46341); *Streptococcus salivarius* (*Strep. sal.*, Genbank M81756); *Streptococcus mutans* (*Strep. mut.*, Genbank L15191); *Bacillus subtilis* (*Bacill. sub.*, Genbank M98359); *Bacillus stearothermophilus* (*Bacill. stear.*, Genbank U12340); *Staphylococcus carnosus* (*Staph. carn.*, SwissProt P23533); *Mycoplasma capricolum* (*Myc. capricol.*, Genbank U15110); *Mycoplasma genitalium* (*Myc. genital.*, TIGR database MG429); *Alcaligenes eutrophus* (*Alcal. eutr.*, SwissProt P23536). Protein sequences were aligned using the PILEUP program from the Wisconsin package [26]. The DNADRAW program (written by M Shapiro of the Division of Computer Research and Technology, NIH) was used for highlighting conserved residues. The numbering above the aligned sequences corresponds only to the residue in the alignment rather than to residues in any of the aligned protein sequences. Residues that are totally conserved are shown as white letters on a black background. Residues that are conservatively retained are shaded gray. The secondary structure assignments are shown above the sequence.

Figure 2


1 70
β1 α1
E.coli. .....MI SGILASPGIA FGKALLLKED EIVIDRKKIS ADQVDQEVER FLSGRAKASA QLETIKTKAG
Rhod.capsul. .....EGRTL VGISSSPGYA LAPVFRFARD EVVFDTDAAD ...AAFETDR LDTALQTAWH ELEELHDEVW
Xanth.camp. .....AI VGIGASPGVA IGIVHRLRAA QTEVADQPIG ...LGDGGVL LHDALTRTRQ QLAAIQDDTQ
Sal.typh. .....MI SGILASPGIA FGKALLLKED EIVIDRKKIS ADKVDQEVER FLSGRAKASA QLEAIKTKAG
Hem.influ. .....MI SGILASPGIA FGKALVLKEE KIVLDTQKIT DDQIDAEVAR FYEGRNAAVE QLNSIRERAL
Strep.sal. .....MTEML KGIAASDGVA VAKAYLLVQP DLSFETVTVE D..TSAEEAR LDAALAASQD ELSVIREKAV
Strep.mut. .....MTEML KGIAASDGVA VAKAYLLVQP DLTFETVSVT D..TQAEEAR LDAALEASQN ELSLIRQKAV
Bacill.sub. ......MQEL KGIGASAGIA IAKAYRLEEP DLTVEKKNIS D..SEAEVSR FDEAIARSKE ELEKIKEHAL
Bacill.stear. MGNAIREKTI HGIAASSGIA IAKAYRLETP DLAAEKRTVA D..VEAEIAR LEAAVAKAKE ELEAIKQHAL
Staph.carn. .....MAKQI KGIAASDGVA IAKAYLLVEP DLSFDNESVT D..TDAEVAK FNGALNKSKV ELTKIRNNAE
Myc.capricol. .....MSKQI KGIAASEGIS LARALVIKET KLDIQKQLIS D..VDQEIIK LEQAIEKSIA DLKKIQQITL
Myc.genital. ......MKKI IGIGVSDGIA VAKAFIIQTP QFDVKKYTNV KMTPTQAKKL LSSAFQKAKK DLEEIKTITV
Alcal.eutr. .....MPFAL HGIPVSRGVA IGRAHLLAPA ALDVSHYLVD EDQLDAEVER LRAARAAVRA ELAALKRDLP
71 140
α2 α3 α4
E.coli. ETFGEEKEAI FEGHIMLLED EELEQEIIAL IKDKHMTADA AAHEVIEGQA SALEEL.DDE YLKERA.ADV
Rhod.capsul. KTSGPARAAI FRAHQEFLHD PEMVAEAKAL IGQGR.SAGF AWHRVFSDRA DMLGAM.KDA VLSGRA.IDL
Xanth.camp. RRLGASDAAI FKAQAELLND TDLITRTCQL MVEGH.GVAW SWHQAVEQIA SGLAAL.GNP VLAGRA.ADL
Sal.typh. ETFGEEKEAI FEGHIMLLED EELEQEIIAL IKDKHMTADA AAHEVIEGQA TALEEL.DDE YLKERA.ADV
Hem.influ. ISLGEEKAAI FEGHLMILED EELEEEILDY LRSNKVNAGV AASKILDQQV TMLSEI.DDE YLKERA.GDI
Strep.sal. ESLGEEAAAV FDAHLMVLAD PEMTGQIKET IRAKQVNAEA ALTEVTDMFI AIFEGMEDNP YMQERA.ADI
Strep.mut. DTLGEEAAAV FDAHLMVLAD PEMIGQIKET IRTKEVNAES ALKEVTDMFV TLFENMEDNP YMQERA.ADI
Bacill.sub. KELGQ..... ....PVKEK ISTDSVNAEF ALKETSSMFV TMFESM.DNE YMKERA.ADI
Bacill.stear. EKLGEDKAAI FAAHLLVLDD PELLNPIKEK IQTERVNAEY ALDETALFFI SMFEAM.DNE YMKERA.ADI
Staph.carn. KQLGADKAAI FDAHLLVLED PELIQPIEDK IKNESVNAAQ ALTDVSNQFI TIFESM.DNE YIAERAKADI
Myc.capricol. KKLGEEKAAI FDAHQDIAND PAIKEEVVEL IKKEKVNAEY ALFTVSNNYF EMFSQLED.P YFKERS.ADI
Myc.genital. KNINQEAGMI FDAHIQILND PTITEQLEQQ L.NKNIHPVI AVDNVFQQTA LMFSEM.DDK YFKERA.SDI
Alcal.eutr. RDAPEELGAF LDVHAMILDD EALAREPEAL IRGRRYNAEW ALTTRLEELM HQFDEI.EDE YLRER.KTDI
141 210
β2 α5 β3
E.coli. RDIGKRLLRN ILGLKIIDLS AIQD..... ...EVILVAA DLTPSETAQL NLKKVLGFIT DAGGRTSHTS
Rhod.capsul. RDAGQRVLQH LGRVRTGE.. T.......HL PTAPCILLAD DLTPSDTARL DPALVRGLAT AQGGPTSHTS
Xanth.camp. RDVGRRVLAQ LDAAAAGAGL T.......DL PEQPCILLAG DLSPSDTANL DTDCVLGLAT AQGGPTSHTA
Sal.typh. RDIGKRLLRN ILGLAIIDLS AIQE..... ...EVILVAA DLTPSETAQL NLQKVLGFIT DAGGRTSHTS
Hem.influ. RDIANRLVKN ILGMYIVDLG DIQE..... ...ESILVAY DLTPSETAQL NLEKVLGVVT DIGGRTSHTS
Strep.sal. RDVTKRVLAN LLGKKLPNPA TINE..... ...ESIVVAH DLTPSDTAQL NKKYVKAFVT NIGGRTSHSA
Strep.mut. RDVAKRVLAH LLGVELPNPA TISE..... ...ESIVIAH DLTPSDTAQL DANYVKAFVT NIGGRTSHSA
Bacill.sub. RDVTKRVTGH LLGVEIPNPS MISE..... ...EVIIVAE DLTPSDTAQL NREFVKGFTT DIGGRTSHSA
Bacill.stear. RDVTKRVLAH LLGVTISNPS LISE..... ...EVVIIAE DLTPSDTAQL NRQYVKGFAT DIGGRTSHSA
Staph.carn. RDVSKRVLAH ILGVELPNPS IVDE..... ...SVVIIGN DLTPSDTAQL NKEYVQGFVT NIGGRTSHSA
Myc.capricol. KDVSLRIISH ILGLEIHDLS TIDK..... ...EVIIISD DLTPSQTAQL DKKFVKGFLT NVGGRTSHAA
Myc.genital. LDLHQRLLSY LTGVKLNDLI RIKS..... ...DVIIVAN DLTPSQTATL NKKYVKGFLT ESGGKTSHAA
Alcal.eutr. RQVVERILKA LAGAPVLVPA PVPALAADGE AATGVIVVAH DIAPADMLQF RHTVFHGFVT DMGGRTSHTA
211 280
α6 β4 α7 β5 β6 α8
E.coli. IMARSLELPA IVGTGSVTSQ VKNDDYLILD AVNNQVYVNP TNEVIDKMRA VQEQVASEKA ELAKLKDLPA
Rhod.capsul. IIARALDIPA VAGVGPRLLD LATGTPVLLD GGAGVIVVAP TEADKARAET AMAALTAQRE LEARERYKPA
Xanth.camp. ILSRTLGLPA LVAAGGQLLD IEDGVTAIID GSSGRLYINP SEQDLDAART HIAEQQAIRE REAAQRALPA
Sal.typh. IMARSLELPA IVGTGSVTAQ VKNGDYLILD AVNNQVYVNP TNDVIEQLRA VQEQVATEKA ELAKLKDLPA
Hem.influ. IMARSLELPA IVGTNKVTKL VNTGDYLILD AINNQVYINP TASQIDELKA LEAKISEEKA ELAKLKDLPA
Strep.sal. IMARTLEIAA VLGTNNITEL VKDGDILAVS GITGEVVINP TEEQIAEFKA AGEAYAKQKA EWALLKDAQT
Strep.mut. IMARTLEIAA VLGTNDITER VKNGDIVAVN GITGQVIINP TEDQIAEFKA AGETYAKQKA EWALLKDAET
Bacill.sub. IMARSLEIPA VVGTKAATGT IQNGVTVIVD GINGDVIIDP SAETVKEYEE KHNAYLAQKA EWAKLVNEPT
Bacill.stear. IMARSLEIPA VVGTKTVTAE VKNGDIVIVD GLDGQVIINP SPELLAQYEQ KRARYEAQKA EWAKLVHEAT
Staph.carn. IMSRSLEIPA VVGTKSITEE VEAGDTIVVD GMTGDVLINP SDEVIAEYQE KRENFFKDKQ ELQKLRDAES
Myc.capricol. IMARSLEIPA ILGLKNITEL VKTDDLIALD GSSGIVELDL NDDDIKNYQT KVQQYIELKE QLKKFKDEPS
Myc.genital. IMARSMEIPA IVGLKNITSK VEDGKTVGIN GRKGIVGFDF SSKDITQWKQ EKELESNFQN ELKQYTNKLV
Alcal.eutr. IVARSLDIPA AVGVQSASEL IRQDDWIIID GDAGLVIVDP TAIILEEYRH RQSERALEKK RLQRLRHTPA

missing in *B. subtilis*, so that it is not clear whether this kink has a functional role.

The α/β subdomain consists of a β sandwich formed by a four-stranded parallel sheet (β1–4) and a three-stranded antiparallel sheet (β1,5 and 6), together with three short helices (α5–7). There is a large bend in the N-terminal β strand (β1) such that it forms the 1st strand of both sheets. The β strands of the two sheets are oriented perpendicular to one another. The β strands of the parallel sheet are connected by helices using the right-handed βαβ motif. The α subdomain is connected to the 1st and 2nd strands (β1 and β2) of the parallel sheet through two extended loops.

The junction between the two subdomains forms a large cleft, in the region of the active-site histidine, running across the middle of the molecule (see Fig. 3, blue arrow). This cleft is divided by helix α4 into two interconnected depressions. The two extended loops, together with the surfaces of the two subdomains on the same region of the cleft, form a large, deep depression (left yellow arrow, Fig. 3). Another smaller, shallow depression (right yellow arrow, Fig. 3) is defined by three helical turns at the N termini of both helices α2 and α4, together with the short helices α5 and α6. These depressions represent regions on the surface of the molecule which may interact with HPr and/or the PEP-binding domain of the protein (EIC).

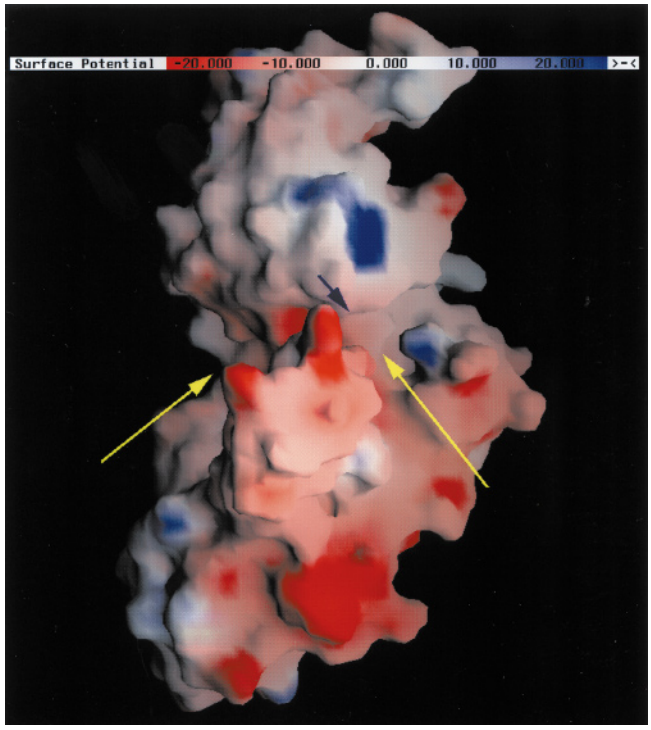
Figure 2 shows a comparison of EIN sequences from a variety of bacteria. These proteins exhibit a high level of sequence conservation in the regions of well defined secondary structure of EIN from *E. coli* (except for α7 and α8). The high level of conservation suggests that all these EI proteins adopt a similar structure. The two regions of greatest sequence conservation are between residues 12–22 and 203–223 (corresponding to the numbering used in the alignment Fig. 2). These residues correspond to positions 4–14 and 184–204 of the *E. coli* sequence and include the site of phosphorylation. These two conserved sequences are found in the same region of the protein (Fig. 1). *B. subtilis* EI is an outlier, having a deletion of 20 residues that would presumably decrease the size of α2, or even eliminate some of the secondary structure in this region. This deletion in *B. subtilis* would cause a significant change in the environment of the active site. The length of the loop region between the α4 helix and β2 strand varies among the different EIs. There are also some other minor sequence differences between EIs that could be readily accommodated without significant changes in the structure.

The active site

The dimerization of intact EI is required for the autophosphorylation by PEP but not for the transfer of the phosphoryl group to HPr [4,7]. EIN only catalyzes the reversible phosphorylation of P–HPr; E1N exists and functions in monomeric form. The phosphorylation site (His189) sits at the N terminus of helix α6 which is connected to the 3rd and 4th β strands. The site is located in the junction of the two subdomains at the border between the two depressions. The Nε atom of His189 makes a hydrogen bond with the Oγ atom of Thr168 (Fig. 4a). In the phosphorylated form of the enzyme, the phosphoryl group is attached to this Nε atom, necessitating a different conformation from that observed here.

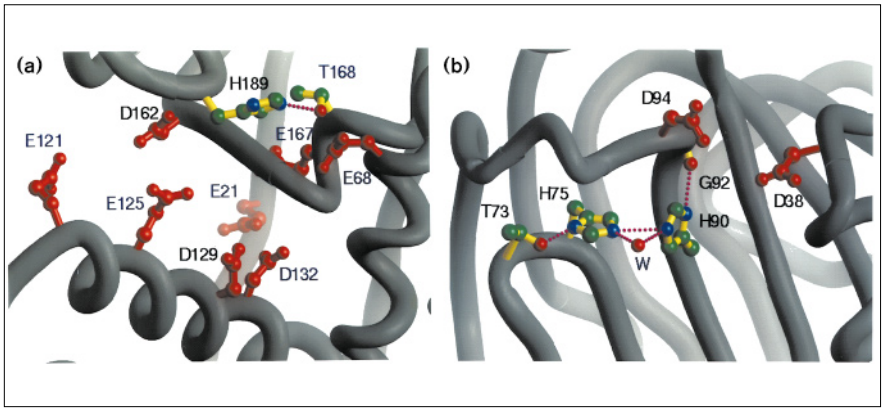
Modeling of the His189 conformation based on the crystal structure shows that the range of allowed χ1 angles, in which the imidazole ring does not make bad contacts with the rest of the protein molecule, includes a gauche plus conformation (χ1 = $^{-}$90 to $^{-}$30 degrees) and a part of the trans conformation (χ1 = 180–210 degrees). Changing the His189 side chain to these conformations would position the imidazole group within the cleft, in either the deep or shallow depression, with most of the allowed conformations lying within the shallow depression. In both of these positions, χ2 can be adjusted to make the Nε of the imidazole ring completely solvent accessible. In addition, with a minor adjustment of the surrounding side-chain atoms, His189,

Figure 3



A GRASP representation of the surface of EIN. The charge potential is colored red for negative and blue for positive. The blue arrow shows the position of His189 and the yellow arrows point to the two depressions in the interdomain region of the protein. The arrow on the left indicates the position of the deep depression, the one on the right the shallow depression.

Figure 4



The active centers of EIN and EIIA^{glc}. (a) The active-site region of EIN showing the high concentration of carboxylate side chains in the vicinity of His189. The two invariant aspartic acid residues (Asp129 and Asp162) are labeled in black. The orientation of His189 (ψ1=168° in one monomer and 166° in the other) permits the formation of a hydrogen bond with Thr168. Hydrogen bonds are represented by magenta dotted lines. (b) The active-site region of EIIA^{glc} from *B. subtilis* showing the phosphorylation site His90 and the invariant aspartate residues (Asp38 and Asp94). In this structure, the Nϵ of His90 points outwards, making a hydrogen bond with a solvent molecule (labeled W).

whether phosphorylated or not, is able to move from one depression to the other and adopt many low energy conformations without making any unfavorable contacts. However, none of the conformations can position the phosphoryl group in such a way that the phosphoryl oxygens can be hydrogen bonded to main-chain nitrogen atoms, as shown in the structural studies of EIIA^{glc} and HPr [8,9].

Comparison of the active site environment of EIN and EIIA^{glc}

EIIA^{glc}, which has no sequence homology or structural similarity to EIN, also interacts with HPr. It accepts the phosphoryl group from P–HPr on the Nϵ of a histidine residue. The phosphoryl group is subsequently transferred to the IIB domain of a permease. As both EIN and EIIA^{glc} interact with HPr, a comparison of their active sites could provide some insight into the nature of the EIN–HPr interaction. In contrast to the numerous possible conformations of His189 in EIN, there is but a single conformation of the site of phosphorylation (His90) in EIIA^{glc} (Fig. 4b), with the Nϵ atom exposed to solvent [8,10].

His90 of EIIA^{glc} is surrounded by a large hydrophobic patch which has been proposed to participate in the HPr–IIA^{glc} interaction [8,11]. The interfaces of HPr–EIIA^{glc} from *B. subtilis* and HPr–EIIA^{mtl} (specific for mannitol transport) from *E. coli* have been mapped using solution NMR studies [12,13]. These studies also showed the involvement of this hydrophobic patch in complex formation. In addition, there is a pair of invariant aspartic acid residues (Asp38 and Asp94 of the *B. subtilis* EIIA^{glc}) which are thought to be important for recognition by HPr. The carboxylate moieties of these two aspartate residues are 4.4 Å apart and about 7 Å from His90 (Fig. 4b). Both structural and modeling studies suggested that this pair of aspartic acid residues is important for the recognition of EIIA^{glc} by HPr through electrostatic interaction with the invariant Arg17 of HPr [8,11]. Mutagenesis studies have shown that this invariant Arg17 is essential for the phosphoryl transfer activity of HPr [14]. In the EIN structure there is no large hydrophobic patch and the active-site cleft in the vicinity of the phosphorylation site histidine contains many charged residues (Fig. 4a). There is a row of four negatively charged residues (Glu121, Glu125, Asp129 and Asp132) on the side of α4 facing the active-site histidine; these four residues are accessible from either the deep or the shallow depression. Other negatively charged residues in the vicinity of His189 are Glu68 (in the shallow depression) and Glu21, Asp162 and Glu167 (in the deep depression). Among these residues, only Asp129 and Asp162 are strictly conserved in all EIs. None of the negatively charged residues mentioned above, except Glu167, form any salt bridges or hydrogen bonds to adjacent atoms of the protein. Unlike EIIA^{glc}, the two invariant Asp residues of EIN are quite far apart with a closest distance between the two carboxyl groups of 10.1 Å (Fig. 4a). However, the clustering of negatively charged residues does suggest that Arg17 in HPr could interact with one or more of them in the vicinity of the active site to strengthen the binding between EIN and HPr. The surface electrostatic potential in both depressions and in the α subdomain is particularly negative (Fig. 3).

These observations suggest that the nature of the interface between the complex of HPr and EIN formed during the phosphotransfer reaction might be quite different from that in the HPr – EIIA^{glc} complex. However, in both cases, the invariant Arg17 of HPr is likely to play an important role in the protein–protein interactions by forming salt bridges with one or more negatively charged side chains from EIN or EIIA^{glc}.

Similarity of EI and pyruvate phosphate dikinase

Pyruvate phosphate dikinase (PPDK) catalyzes a reaction that has some similarities to the EI reaction. The overall reaction catalyzed by PPDK (PEP + AMP + PPi $\leftrightarrow$ pyruvate + ATP + Pi) occurs by a mechanism in which a histidine residue is phosphorylated by PEP [15]. PPDK consists of three domains: an ATP/Pi-binding domain; a

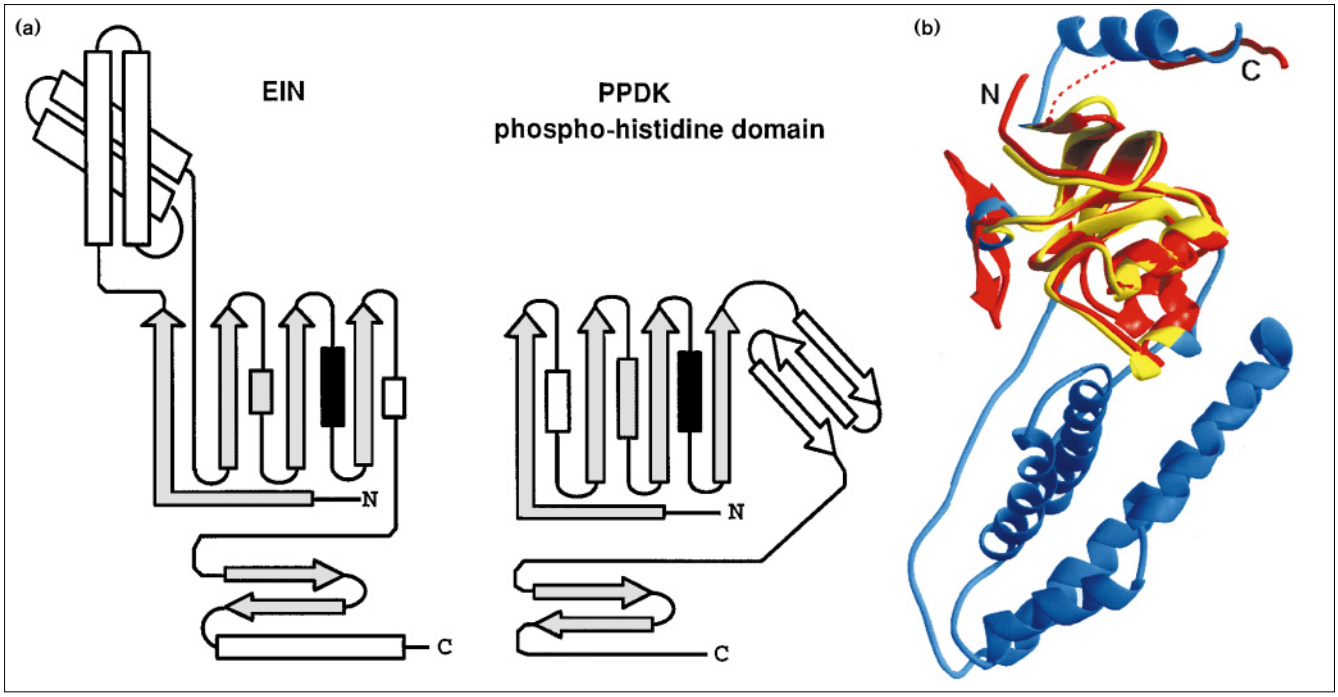
phosphohistidine domain; and a PEP/pyruvate-binding domain. The crystal structure of PPDK from *Clostridium symbiosum*, with all three domains intact, has been recently determined [16].

The sequences of two regions of EI are highly homologous to corresponding regions in PPDK [15]. One of these regions is located in the putative PEP-binding site of EI. In this region there is 36% sequence identity (for 201 residues in the C-terminal domain of EI) when compared to the PEP/pyruvate-binding domain of PPDK. This sequence homology strongly suggests that the structure of the C-terminal domain of EI will, like PPDK, also be an eightfold α/β barrel; by analogy with PPDK the PEP-binding site would be located at the center of the C-terminal end of the barrel. The second region of high homology with PPDK occurs in a sequence of 12 residues in EIN which contain the phosphorylation site histidine. The remainder of the sequences show little similarity between the two proteins. As both EIN and the phosphohistidine domain of PPDK can be phosphorylated on a histidine residue by PEP, it has been speculated that, despite the lack of overall sequence homology, their tertiary structures may be similar (O Herzberg, personal communication).

A structural comparison of the domain containing the phosphorylation site His455 in PPDK (residues 390–504) and EIN indicates that they indeed have very similar folds (Fig. 5a). Superposition of the two corresponding regions (Fig. 5b) shows the similarity of the three-dimensional structures. The root mean square (rms) difference between the main-chain Cα atoms of the two structures is 1.36 Å for 81 target pairs and twenty-three of these pairs have identical amino acids. The majority of the overlapping pairs are in the β sheet region but only two of the three helices align well (Fig. 5a). One of these is the active-site helix (shown in black) containing the phosphorylation site histidine. The second is the helix (shown in grey) connecting the 2nd and 3rd strands of the parallel sheet, although this helix is shorter in EIN than in PPDK. Both helices are located in the shallow part of the cleft (Fig. 5b).

The strands of the β sandwich of both molecules align well and have identical chain connectivities, although there are some insertions and deletions (Fig. 5a). The three antiparallel β strands that connect the two sheets of the β sandwich in PPDK are replaced by a one-turn helix in EIN, located in the deep part of the cleft. The PPDK structure lacks the pair of α hairpins which are replaced by

Figure 5



Comparison between EIN and the phosphohistidine domain of pyruvate phosphate dikinase (PPDK). (a) Schematic diagram of the topology of the two molecules. The helices containing the active-site histidine are shown in black. The regions of topology common to both proteins are in gray. (b) Ribbon drawing of EIN superimposed on the corresponding region of PPDK (red). The regions of EIN containing the residues used for the alignment are in yellow and the remainder of the structure is shown in blue. The red dashed line corresponds to the region in PPDK that does not have clearly defined electron density. The N and C terminus of the region of PPDK shown are indicated by N and C, respectively.

a two-turn α helix. As PPDK does not interact with HPr, this suggests that the α subdomain of EIN may function in HPr binding.

Autophosphorylation of EI

A number of proteins that are autophosphorylated on histidine residues have been shown to exist as dimers. In the case of the *E. coli* chemotaxis protein, CheA (autophosphorylated by ATP), the phosphoryl transfer is intersubunit [17]. The autophosphorylation of nitrogen regulator II by ATP also occurs within the dimer by a *trans*, intersubunit mechanism [18]. PPDK is autophosphorylated by a mechanism analogous to that of EI in that the phosphoryl donor is PEP. In this case, the analysis of the crystal structure indicates that there is no phosphotransfer across monomers [16]. With respect to EI, no experiments have been carried out to evaluate the possibility of a *trans*-phosphorylation in the dimer. The crystal structure reported here provides no insight into the spatial organization of the PEP-binding site in relation to the site of phosphorylation. It is therefore premature to make any conclusions regarding the mechanism of autophosphorylation of EI.

Molecular modeling of the EIN and HPr complex

We have modeled the transition state complex between EIN and HPr using the phosphate linkage between His189 of EIN and His15 of HPr as a tether, assuming there is no large conformational transition associated with the interaction. The flexibility of the side chain of His189 allows for several possible models, even if we impose the condition that the Arg17, which is essential for HPr activity, should make a salt bridge with one or more of the carboxylates that surround the active site of EIN.

Two plausible models for the interaction between HPr and EIN were evaluated. In the first model HPr interacts with the shallow depression of EIN (model I), in the second HPr interacts with the deep depression (model II). These models were achieved by adjusting the χ1 angle of His189 of EIN (Table 1). In both cases, the relative orientations were selected by visual inspection to minimize bad contacts. One common feature in both models is the interaction of Arg17 of HPr and Asp129, one of the invariant residues near His189 of EIN, although the modes of interaction in the two models are quite different.

Energy minimization caused both models to converge to similar energies, although model I is characterized by a smaller contact area and fewer interactions than model II (Table 1). However, model II requires many more changes in the positions of the main-chain atoms of EIN and HPr. These changes result in the loss of some secondary structures in both molecules. When HPr interacts with the shallow depression in EIN no such changes are required and all of the secondary structures are preserved. Accordingly, we favor model I with HPr binding in the shallow depression (Fig. 6a). The surface electrostatic potential of HPr is rather neutral around the active site, except in the vicinity of the invariant Arg17. In the proposed model only two buried charged groups in EIN do not interact with Arg17 of HPr. They are Glu68, which interacts with the main-chain and side-chain nitrogen atoms of Asn12 of HPr, and Asp120 which interacts with Lys27 of HPr.

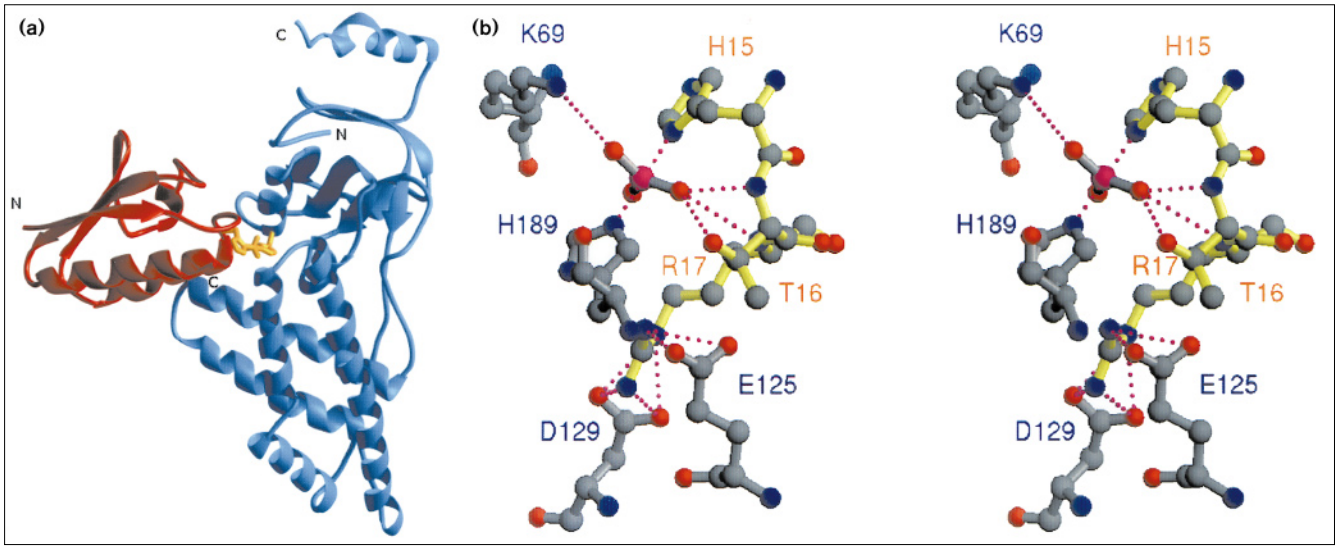
The complex formed in this model involves the interaction of a few side chains with the phosphoryl oxygens; these interactions could stabilize the pentacoordinated phosphorus in the transition state (Fig. 6b). The main-chain nitrogen atoms of HPr residues Thr16 and Arg17 form hydrogen bond contacts with one oxygen of the trigonal phosphoryl group. Other interactions of the phosphoryl oxygens occur with the Nζ of Lys69 of EIN, and the Oγ of Thr16 of HPr, but neither of these are conserved residues. As in both EIN and HPr the phosphorylation

Table 1

Comparison of two proposed models of the EIN–HPr complex.

	Model I		Model II	
	EIN	HPr	EIN	HPr
Rms difference in Cα coordinates before and after energy minimization (Å)	0.74	0.64	0.98	0.99
Side chain (χ1) of phosphorylation site histidine	–155°	55°	–45°	55°
Buried area (Å^2)	684	657	1520	1475
Number of residues in the contact area	16	12	53	29
Initial energy before minimization E_{final} (kcal mol^{-1})	0.18×10^8		0.84×10^8	
Energy after minimization E_{final} (kcal mol^{-1})	–11343.4		–11205.3	
Number of electrostatic interactions	24		57	
Number of hydrophobic interactions	29		85	

Figure 6



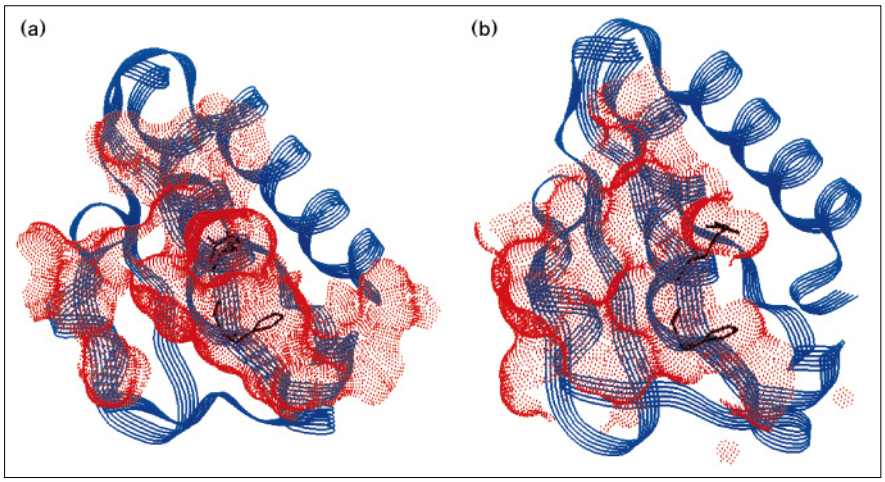
Model of EIN complexed with HPr. (a) The ribbon drawing shows EIN in blue complexed with HPr shown in red. A phosphoryl group linking the two histidine residues is depicted in yellow. (b) Stereo drawing of the active-site residues proposed to be involved in phosphoryl transfer between EIN (gray) and HPr (yellow). The figure illustrates the hydrogen bonds that may stabilize the transition state (shown as magenta dotted lines). Atoms are shown in standard colors with phosphorus in magenta.

site is located at the N terminus of an α helix, it is likely that the helix dipoles also play a role in stabilizing the transition state.

The area of the buried surface and the number of interacting residues in the favored complex are very similar to those proposed by Herzberg [11] in a model of the *B. subtilis* EIIA^{glc}–HPr complex. Figures 7a and b show the interacting surfaces for the two HPrs (*E. coli* and *B. subtilis*) complexed to EIN and EIIA^{glc}, respectively. Not surprisingly, as the common residues His15 and Arg17 form the focus of the interaction, these two surfaces share a number of common features. However, away from the central region, different parts of the protein surfaces are involved in the interaction due to the differences in structure and the active-site environment of EIN and EIIA^{glc}.

This modeling study suggests that it is possible to effect phosphoryl exchange between EIN and HPr without the need for major structural changes. This is consistent with the suggestion by Herzberg, in modeling the EIIA^{glc}–HPr complex that, in general, phosphotransfer reactions occur without major conformational transitions [11].

Figure 7



The surfaces of HPr that contact EIN and EIIA^{glc}. The figure shows the interface of the (a) *E. coli* HPr–EIN complex and (b) the *B. subtilis* HPr–EIIA^{glc} complex; the interface is shown as red dots. The ribbon structure of the HPrs are shown in blue. Arg17 and His15, are depicted as black stick models.

Biological implications

The phosphoenolpyruvate: sugar phosphotransferase system (PTS) is an important pathway for bacterial sugar transport. EI, the first protein in the pathway, is autophosphorylated by phosphoenolpyruvate (PEP) and then transfers its phosphoryl group to the phosphocarrier protein HPr. It is known that EI has a protease-sensitive C-terminal domain (EIC) and a protease-resistant N-terminal domain (EIN) [4]. EIN appears to not undergo the dimerization reaction required for the PEP-dependent autophosphorylation of EI [4], suggesting that EIC is necessary for that process. Although EIN is incapable of PEP-dependent autophosphorylation [4], it can participate in a reversible phosphorylation reaction with P–HPr [4,5]. EIN can effectively transfer its phosphoryl group to HPrs from many different species. However, *E. coli* EI only permits phosphoryl transfer to *E. coli* HPr [5], implying that the interaction of EI with HPr must involve some participation of EIC.

The crystal structure of EIN presented here provides the first insight into the fold of EIN. It establishes clearly that the **α/β** subdomain resembles the phosphohistidine domain of pyruvate phosphate dikinase (PPDK), which is capable of an analogous PEP-dependent autophosphorylation reaction [15]. The finding that the **α** subdomain of EIN has no counterpart in PPDK, a protein which does not interact with HPr, provides a strong argument that this subdomain is important in HPr recognition.

The determination of the EIN crystal structure has allowed us to model the interaction of EIN with HPr. The possibility for complex formation without major main-chain conformational transitions of EIN is demonstrated. The proposed phosphoryl transfer transition state complex contains the interaction of two invariant residues, Arg17 of HPr and Asp129 of EIN. This proposal is supported by previous mutagenesis studies which indicated that Arg17 is essential for the phosphoryl transfer activity of HPr [14]. In summary, this study provides the first clues concerning the mechanism of protein–protein interaction in the first reaction of the PTS pathway, and provides a basis for further mutagenic analysis.

Materials and methods

Crystallization

Single crystals of EIN were obtained at room temperature (~20°C) by vapor diffusion in hanging drops. Protein drops were equilibrated against reservoir solutions containing 20–25% (w/v) PEG 4000, 100 mM cacodylate (pH 6.0) and 50 mM potassium sulfate. The drops consisted of protein at a concentration of 6–12 mg ml^{-1} and 10 mM Tris buffer (pH 7.4), diluted with an equal volume of reservoir solution. Crystals in the shape of thin plates appeared after 1–4 days and some of them were single. The space group of the crystals is $P2_12_12_1$. The crystals were very fragile and thin layers usually fell out when they were moved. The crystals used for data collection were usually 0.05–0.1 mm thick and had various degrees of disorder in a direction perpendicular to the plate surface, corresponding to the b axis. Three native data sets were collected (native 1, native 2 and native 3). Although the crystals used for data collection had the same size and morphology, their diffraction quality varied from 2.5–3.5 Å resolution (Table 2). The cell dimensions of these crystals are very similar: *a* = 47.42 Å, *b* = 75.78 Å, *c* = 170.28 Å (for native 1); *a* = 47.60 Å, *b* = 76.30 Å, *c* = 170.50 Å (for native 2); and *a* = 47.54 Å, *b* = 76.38 Å, *c* = 171.01 Å (for native 3). However, a comparison of these data sets in pairs yielded R values of approximately 20%, indicating that they are non-isomorphous. We attribute the problem of non-isomorphism to the weak lattice interactions along the b axis that could result in small differences in crystal packing (see note [*] in Table 2). There are two EIN monomers in the asymmetric unit, and the solvent content of the crystal is ~50% by volume.

Structure determination

The structure was determined at 3.0 Å resolution by multiple isomorphous replacement (MIR). The X-ray diffraction data were collected at room temperature (~20°C) on an RAXIS-II imaging plate system with CuK_α X-rays from a Rigaku rotating anode generator. The data statistics of the three non-isomorphous native data sets and those derivatives used for structure determination are listed in Table 2. The data were processed using the RAXIS-IIC data processing software package or the program DENZO [19]. Comparison with the three native data sets indicated that only native 2 was isomorphous with the two heavy-atom derivatives and this was used for the initial MIR phase determination. The computer program package PHASES (W Furey & S Swaminathan, 14th American Crystallographic Association Meeting, April 8–13, 1990, New Orleans, Abstr. PA33) was used for heavy-atom parameter refinement, MIR phase calculation, solvent-flattening and non-crystallographic two-fold map averaging. The native Patterson map has four symmetry related 28.0σ peaks in the y = 0.5 Harker section. The location of the peaks indicates that the non-crystallographic two-fold axis is parallel to the crystallographic y axis and it intersects (0.0, 0.0, 0.66), in fractional coordinates, based on the chosen origin. All the heavy-atom binding sites also followed this non-crystallographic symmetry. After four cycles of heavy-atom parameter refinement and solvent-flattening, a phase set with the mean figure of merit of 0.49 was derived. Anomalous difference signals from both derivatives were included in phase calculations. This map showed clearly that the protein contains two very long α hairpins and a domain containing two β sheets. Sixteen cycles of map averaging and solvent-flattening improved the α helical region of the map. The quality of the electron-density map in this region enabled unambiguous tracing of the polypeptide chain and the assignment of sequence designation to each residue. The electron-density map in the the β sheet region has rather poor connectivity but the orientation of the β sheets are very clear in both monomers. The map also showed that the α helical regions are tightly packed in the crystal lattice and that not many crystal contacts can be found in the β sheet region. Map interpretation and model fitting were done by using the computer graphics program O [20]. A monomer containing 229 amino acid residues was first constructed and the model for the second molecule in the asymmetric unit was generated by applying the non-crystallographic two-fold transformation to the first molecule.

Refinement of the structure

The native data set native 1 which has higher resolution and completeness was used for the refinement of this model. One cycle of rigid-body refinement against native 1 data was done on the initial model, before further refinement. 10% of the data have been left out of the refinement. Using these reflections, a free R factor was calculated at each step of the refinement to avoid over-fitting the data [21]. The molecular dynamics refinement program X-PLOR [22] was used to refine the structure. Eight cycles of refinement were carried out. Each cycle of refinement consists of one round of simulated annealing using the slow-cooling protocol of X-PLOR, restrained individual B factor refinement and manual model fixing on each monomer. The non-crystallographic two-fold (NCS) restraint was applied to both main-chain and side-chain atoms for the first three cycles of refinement. At the third

Table 2

Crystallographic statistics.

Data set*	Resolution (Å)	Number of observations	Unique reflections	Completeness†	R_{merge}‡ (%)	Number of sites	R_{iso}‡ (%)	R_{cen}‡ (%)	$<F_H>/<\varepsilon>$‡
Native 1	2.5	128 042	20 342	0.92 (0.66)	9.7	–	(20.1)#	–	–
Native 2	2.7	55 594	14 957	0.84 (0.70)	9.4	–	(23.4)§	–	–
Native 3	3.0	80 720	12 667	0.96 (0.91)	13.9	–	(20.1)¶	–	–
EINSa**	3.0	55 099	10 144	0.78 (0.63)	9.5	2	37.6‡‡	60.0	1.65
EINPb††	3.0	89 348	12 128	0.93 (0.82)	11.9	6	34.0‡‡	57.0	1.64

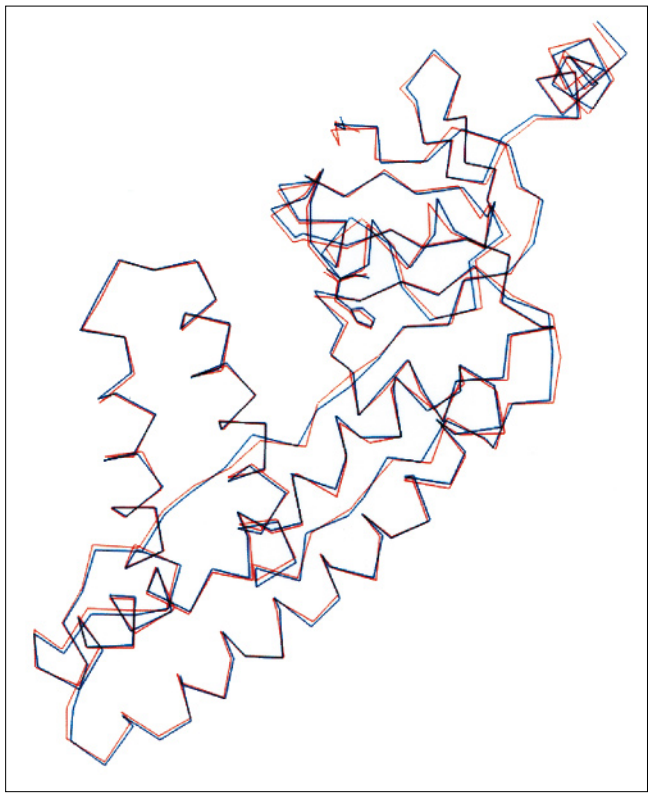
*All the soaking experiments were carried out in 25% PEG 4K, 100 mM MES buffer (pH 6.2) and 50 mM K_2SO_4. Crystals used for collecting native 2 and native 3 data sets were treated under the same conditions. The crystal used for collecting the native 1 data set was mounted directly from the crystallization drop. †Numbers in parentheses are the completeness of the data in the highest resolution shell: native 1 (2.55–2.50 Å); native 2 (2.76–2.70 Å); native 3, EINSa and EINPb (3.07–3.00 Å). ‡$R_{merge}=\Sigma_h\Sigma_i|<I>-I(h)i|/\Sigma_h\Sigma_i I(h)_i$. $R_{iso}=\Sigma_h||F_{PH}obs|-|F_Pobs||/\Sigma_h|F_Pobs|$. $R_{cen}=\Sigma_{h\ centric}||F_{PH}obs|-|F_Pobs+F_Hcalc||/\Sigma_{h\ centric}||F_{PH}obs|-|F_Pobs||$, where h = Miller indices; obs = observed; calc = calculated; F_P, F_{PH}, F_H = native protein, derivative protein, and heavy-atom structure factors, respectively. $<\varepsilon>$, mean lack of closure error; $<F_H>$, mean heavy-atom scattering. #R_{iso} of native 1 and native 2. §R_{iso} of native 2 and native 3. ¶R_{iso} of native 1 and native 3. **2.5 mM samarium acetate, 5.5 day soak. ††10 mM trimethyl lead acetate, 3 day soak. ‡‡R_{iso} between derivative data and native 2.

and fourth cycles, this restraint was applied only to the main-chain atoms; two flexible regions (residues 19–23 and 142–155) were released from the restraint. For the last three cycles of refinement the NCS restraint was released. Water molecules were incorporated into the model at the last two cycles. Application of the NCS restraint, especially at the initial stage of the refinement while there were still errors in the model, is crucial to avoid over-fitting the data. However, the restraint was completely released at the fifth cycle of refinement, resulting in small decreases in both the free R factor (~0.2%) and working R factor (~0.4%). The rms difference in the main-chain position of the two molecules is 0.18 Å before the release of the NCS restraint and is 0.51 Å in the final model, indicating that there are no major differences in conformation (Fig. 8). The final model consists of 247 residues in one molecule (monomer A), 248 residues in the other (monomer B) and 64 water molecules. The working R factor of the model is 20.5% and the free R factor is 30.6% for the 16 764 reflections between 10.0 and 2.5 Å for which $F \geq 3\sigma(F)$. The rms deviations from ideal values for bond lengths and bond angles are 0.014 Å and 1.9°, respectively. The average rms differences in B values for 495 main-chain residues are 2.4 $Å^2$ and 3.7 $Å^2$ for 409 side-chain residues. Figure 9 shows the current electron-density map in the α helical region. The quality of the final map was further assessed by calculating the real-space correlation coefficient, using the RS_fit function in the program O, to evaluate how well the model fits the density [20]. Most of the residues have coefficients between 0.94–0.80. Some of the less well defined regions of the map, such as the two loops that link the two subdomains, are slightly less well correlated (between 0.90–0.66 in monomer A, and 0.90–0.72 in monomer B).

Modeling of the EIN–HPr complex

The coordinates for the refined model of EIN, from this study, and *E. coli* HPr obtained from the Brookhaven Protein Data Bank (entry code 1POH) [23] were used for the modeling of the complex. The software package QUANTA was used for building the phosphohistidine on the Nε atom of His189 of EIN and the docking of the two molecules. The model of the transition state was built based on the trigonal bipyramidal geometry of phosphorus required by the phosphoryl transfer via an associative pathway [24,25]. The phosphorus atom was positioned 2.0 Å away from the Nε atom of His189 of EIN and the Nδ of His15 of HPr, with both nitrogen atoms in apical positions. The HPr molecule was then rotated around the apical axis and the χ1 and χ2 dihedral angles of both histidine side chains were manually adjusted. This was to accommodate the geometry of the transition state and to avoid interpenetration of the two molecules. The possibility that Arg17 might form

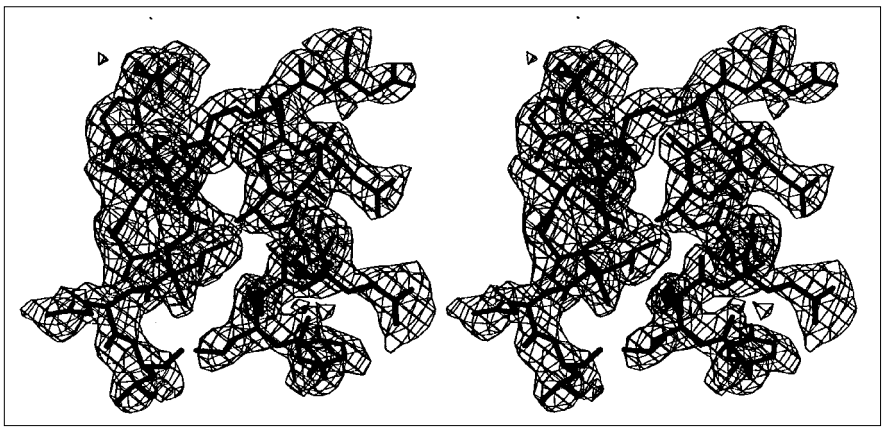
Figure 8



Superposition of the Cα backbone of molecule A, shown in blue, and molecule B, shown in red, of EIN in the asymmetric unit.

Figure 9

Stereo drawing of a representative region of electron density shown in a 2Fo–Fc map. The figure shows a superposed model of the end of helix α3 and the beginning of helix α4.



electrostatic interactions with negatively charged residues in the vicinity of the active-site His189 was taken into account when the orientation was chosen. However, adjustment of the side-chain conformation of His15 of HPr was kept to a minimum. Another assumption used in this modeling exercise was that there is no large conformational transition in either molecule associated with the formation of the complex. Energy minimizations were performed using CHARMM energy minimization available in X-PLOR. Enough steps of conjugate gradient energy minimization were applied until the energy converged.

Accession number

The coordinates for this structure have been deposited in the Brookhaven Protein Data Bank with accession code 1ZYM.

Acknowledgements

We wish to thank Dr Osnat Herzberg for providing us with the coordinates of PPDK in advance of publication. We acknowledge the assistance of Dr Peng-Peng Zhu in the computer analysis of enzyme I sequences.

References

1. Postma, P.W., Lengeler, J.W. & Jacobson, G.R. (1993). Phosphoenolpyruvate: carbohydrate phosphotransferase systems of bacteria. *Microbiol. Revs.* **57**, 543–594.
2. Herzberg, O. & Klevit, R. (1994). Unraveling a bacterial hexose transport pathway. *Curr. Opin. Struct. Biol.* **4**, 814–822.
3. Lee, B.R., Lecchi, P., Pannell, L., Jaffe, H. & Peterkofsky, A. (1994). Identification of the N-terminal domain of enzyme I of the *Escherichia coli* phosphoenolpyruvate:sugar phosphotransferase system produced by proteolytic digestion. *Arch. Biochem. Biophys.* **312**, 121–124.
4. Licalsi, C., Crocenzi, T.S., Freire, E. & Roseman, S. (1991). Sugar transport by the bacterial phosphotransferase system. Structural and thermodynamic domains of enzyme I of *Salmonella typhimurium. J. Biol. Chem.* **266**, 19519–19527.
5. Seok, Y.-J., Lee, B.R., Zhu, P.-P. & Peterkofsky, A. (1996). Importance of the carboxyl-terminal domain of enzyme I of the *Escherichia coli* phosphoenolpyruvate:sugar phosphotransferase system for phosphoryl donor specificity. *Proc. Natl. Acad. Sci. USA* **93**, 347–351.
6. Seok, Y.-J., Lee, B.R., Gazdar, C., Svenson, I., Yadla, N. & Peterkofsky, A. (1996). Importance of the region around glycine-338 for the activity of enzyme I of the *Escherichia coli* phosphoenolpyruvate: sugar phosphotransferase system. *Biochemistry* **35**, 236–242.
7. Chauvin, F., Brand, L. & Roseman, S. (1994). Sugar transport by the bacterial phosphotransferase system. Characterization of the *Escherichia coli* enzyme I monomer/dimer equilibrium by fluorescence anisotropy. *J. Biol. Chem.* **269**, 20263–20269.
8. Liao, D.-I., Kapadia, G., Reddy, P., Saier, M.H., Jr., Reizer, J. & Herzberg, O. (1991). Structure of the IIA domain of the glucose permease of *Bacillus subtilis* at 2.2 Å resolution. *Biochemistry* **30**, 9583–9594.
9. van Nuland, N.A.J., Boelens, R., Scheek, R.M. & Robillard, G.T. (1995). High-resolution structure of the phosphorylated form of the histidine-containing phosphocarrier protein HPr from *Escherichia coli* determined by restrained molecular dynamics from NMR-NOE data. *J. Mol. Biol.* **246**, 180–193.
10. Worthylake, D., Meadow, N.D., Roseman, S., Liao, D.-I., Herzberg, O. & Remington, S.J. (1991). Three-dimensional structure of the *Escherichia coli* phosphocarrier protein III^{glc}. *Proc. Natl. Acad. Sci. USA* **88**, 10382–10386.
11. Herzberg, O. (1992). An atomic model for protein-protein phosphoryl group transfer. *J. Biol. Chem.* **267**, 24819–24823.
12. Chen, Y., Reizer, J., Saier, M.H., Jr., Fairbrother, W.J. & Wright, P.E. (1993). Mapping of the binding interfaces of the proteins of the bacterial phosphotransferase system, HPr and IIA^{glc}. *Biochemistry* **32**, 32–37.
13. van Nuland, N.A.J., Kroon, G.J.A., Dijkstra, K., Wolters, G.K., Scheek, R.M. & Robillard, G.T. (1993). The determination of the IIA^{mtl} binding site on HPr of the *Escherichia coli* phosphoenol-pyruvate dependent phosphotransferase system. *FEBS Lett.* **315**, 11–15.
14. Sharma, S., Georges, F., Delbaere, L.T.J., Lee, J.S., Klevit, R.E. & Waygood, E.B. (1991). Epitope mapping by mutagenesis distinguishes between the two tertiary structures of the histidine-containing protein HPr. *Proc. Natl. Acad. Sci. USA* **88**, 4877–4881.
15. Pocalyko, D.J., Carroll, L.J., Martin, B.M., Babbitt, P.C. & Dunaway-Mariano, D. (1990). Analysis of sequence homologies in plant and bacterial pyruvate phosphate dikinase, enzyme I of the bacterial phosphoenolpyruvate:sugar phosphotransferase system and other PEP-utilizing enzymes. Identification of potential catalytic and regulatory motifs. *Biochemistry* **29**, 10757–10765.
16. Herzberg, O., *et al.*, & Dunaway-Mariano, D. (1996). Swiveling-domain mechanism for enzymatic phosphotransfer between remote reaction sites. *Proc. Natl. Acad. Sci. USA* **93**, 2652–2657.
17. Stock, A.M. & Mowbray, S.L. (1995). Bacterial chemotaxis: a field in motion. *Curr. Opin. Struct. Biol.* **5**, 744–751.
18. Ninfa, E.G., Atkinson, M.R., Kamberov, E.S. & Ninfa, A.J. (1993). Mechanism of autophosphorylation of *Escherichia coli* nitrogen regulator II (NR_{II} or NtrB): trans-phosphorylation between subunits. *J. Bacteriol.* **175**, 7024–7032.
19. Otwinowski, Z. (1993). Oscillation data reduction program. In *Data Collection and Processing.* (Sawyer, L., Isaccs, N. and Bailey, S., eds), pp. 56–62, SERC Daresbury Laboratory, Warrington, UK.
20. Jones, T.A., Zou, J.-Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst.* A **47**, 110–119.
21. Brünger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–475.
22. Brünger, A.T., Kuriyan, J. & Karplus, M. (1987). Crystallographic R factor refinement by molecular dynamics. *Science* **235**, 458–460.

23. Jia, Z., Quail, J.W., Waygood, E.B. & Delbaere, L.T.J. (1993). The 2.0 Å resolution structure of *Escherichia coli* histidine-containing phosphocarrier protein HPr. A redetermination. *J. Biol. Chem*. 268, 22490–22501.
24. Begley, G.S., Hansen, D.E., Jacobson, G.R. & Knowles, J.R. (1982). Stereochemical course of the reactions catalyzed by the bacterial phosphoenolpyruvate: glucose phosphotransferase system. *Biochemistry* 21, 5552–5556.
25. Knowles, J. R. (1980). Enzyme-catalyzed phosphoryl transfer reactions. *Annu. Rev. Biochem*. 49, 877–919.
26. Devereux, J., Haeberli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res*. 12, 387–395.